

SPRING 2016

**NEW YORK UNIVERSITY SCHOOL OF LAW
COLLOQUIUM ON TAX POLICY AND
PUBLIC FINANCE**

“Preference-Shifting and the Non-Falsifiability
of Optimal Tax Theory”

Professor Theodore P. Seto
Loyola Law School

March 8, 2016
Vanderbilt-208
Time: 4:00-5:50 pm
Number 7

SCHEDULE FOR 2016 NYU TAX POLICY COLLOQUIUM
(All sessions meet on Tuesdays from 4-5:50 pm in Vanderbilt 208, NYU Law School)

1. January 19 – Eric Talley, Columbia Law School. “Corporate Inversions and the unbundling of Regulatory Competition.”
2. January 26 – Michael Simkovic, Seton Hall Law School. “The Knowledge Tax.”
3. February 2 – Lucy Martin, University of North Carolina at Chapel Hill, Department of Political Science. “The Structure of American Income Tax Policy Preferences.”
4. February 9 – Donald Marron, Urban Institute. “Should Governments Tax Unhealthy Foods and Drinks?”
5. February 23 – Reuven S. Avi-Yonah, University of Michigan Law School. “Evaluating BEPS”
6. March 1 – Kevin Markle, University of Iowa Business School. “The Effect of Financial Constraints on Income Shifting by U.S. Multinationals.”
7. **March 8 – Theodore P. Seto, Loyola Law School, Los Angeles. “Preference-Shifting and the Non-Falsifiability of Optimal Tax Theory.”**
8. March 22 – James Kwak, University of Connecticut School of Law. “Reducing Inequality With a Retrospective Tax on Capital.”
9. March 29 – Miranda Stewart, Australian National University. “Transnational Tax Law: Reality or Fiction, Future or Now?”
10. April 5 – Richard Prisinzano, U.S. Treasury Department, and Danny Yagan, University of California at Berkeley Economics Department. "Partnerships in the United States: Who Owns Them and How Much Tax Do They Pay?"
11. April 12 – Lily Kahng, Seattle University School of Law. “Who Owns Human Capital?”
12. April 19 – James Alm, Tulane Economics Department, and Jay Soled, Rutgers Business School. “Whither the Tax Gap?”
13. April 26 – Jane Gravelle, Congressional Research Service.
“Policy Options to Address Corporate Profit Shifting: Carrots or Sticks?”
14. May 3 – Monica Prasad, Northwestern University Department of Sociology.
“The Popular Origins of Neoliberalism in the Reagan Tax Cut of 1981.”

Preference-Shifting and
the Non-Falsifiability of Optimal Tax Theory

Theodore P. Seto

Optimal tax theory is based on a core factual assumption – that preferences reflect welfare. In practice, this assumption is neither tested nor questioned. Science requires falsifiability – of both theories and their factual predicates. That the core factual assumption upon which optimal tax theory is based is neither tested nor questioned is therefore problematic.

Advertising offers a useful context in which to think about the extent to which preferences do or do not reflect welfare. Some advertising conveys information. In the language of market theory, it remedies informational failures. To the extent it does so, market theory would not characterize it as changing preferences; market theory would rather characterize it as allowing pre-existing preferences to be satisfied more efficiently. Not all advertising, however, achieves its goals by conveying information. Some of the most effective campaigns in advertising history have been remarkably devoid of content: “The Ultimate Driving Machine,” “You Are Now Free to Move About the Country,” “The Best Part of Waking Up is Folgers in Your Cup!” “Just Do It,” “Where’s the Beef?” “The Real Thing,” “Delta Is Ready When You Are,” “Mmm! Mmm! Good!” “Nobody Doesn’t Like Sara Lee,” “I’d Like to Buy the World a Coke,” “Great Taste, Less Filling,” “Think Small,” “Marlboro Man,” “Does She or Doesn’t She?,” “A Diamond is Forever,” “Zoom zoom,” “Got Milk?” A scientist would likely consider the possibility that advertising changes preferences, perhaps even in ways that do not enhance consumer welfare.

Measuring any resulting wedge between preferences and welfare is extremely difficult, in part because we lack a consensus definition of welfare. Such evidence as there is, however, suggests that a significant such wedge may exist.

If so, the validity and explanatory power of the optimal tax model would be enhanced if the possibility that preferences do not reflect welfare could be generally incorporated. This paper outlines one possible approach to doing so. Specifically, it postulates the existence of “preference-shifting” – where businesses shift the preferences of potential consumers in ways that are not commensurately welfare-enhancing and then satisfy the resulting shifted preferences – and suggests a simple way to incorporate preference-shifting into the standard model.

The paper then focuses on two of the model’s canonical assertions: (1) that taxes produce deadweight loss (Harberger 1964), and (2) that 100 percent of all taxes are borne by human beings, the only question being

which. If preference-shifting exists, the paper demonstrates (1) that a tax less than or equal to the amount of any such preference-shift will not result in any tax deadweight loss and will reduce preference-shifting deadweight loss; as a result, such a tax will increase aggregate social welfare, and (2) that a portion of the welfare impact of the tax up to the amount of the preference-shifting dead-weight loss will be recovered through elimination of that loss and will not be borne by any human being. In other words, whether the assumption that preferences reflect welfare is true matters.

Part I: The Structure of Optimal Tax Theory

The conceptual structure of optimal tax theory is not often explicitly articulated or explored. As commonly applied, the theory is normative, not descriptive: “The theory of optimal taxation has, for the past two decades, been the reigning normative approach to taxation.” (Slemrod 1990) Normative claims are not generally required to be falsifiable. Nor are they part of science.

Philosophically, optimal taxation is a branch of welfarism, which holds that rules are morally justified if and to the extent they maximize some social welfare function – most commonly, the sum of all relevant individual welfare functions. This paper accepts, arguendo, this normative starting point.

At the core of welfarism’s analytic method, however, is a factual claim: that an individual’s preferences correctly reflect his or her welfare. If an individual prefers something, the theory assumes, satisfaction of that preference will enhance his or her welfare. The extent to which his or her welfare will be increased can be measured by the amount he or she is willing to pay for it – an observable behavior.

This assumption – that preferences reflect welfare – allows welfarism to imbue the descriptive tools of microeconomics with moral force. Consider standard supply and demand analysis. Microeconomics asserts that, in general, supply and demand adjust to the point at which they are in equilibrium – a descriptive claim with no inherent moral valence. If observed choices reflect preferences which in turn reflect welfare, however, then (1) the height of the demand curve at any quantity equals the marginal social benefit of consumption of that quantity, (2) the height of the supply curve equals the marginal social cost of production of that quantity, and (3) total surplus provides a measure of the net increase in social welfare by reason of market exchanges at the equilibrium price and quantity, all costs and benefits having been internalized. In the absence of market failure, supply and demand adjust automatically to maximize social welfare. In the absence of market failure, therefore, we *should* respect free markets.

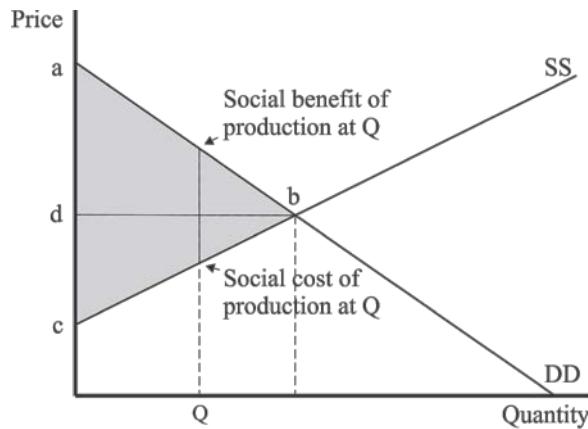


Figure 1

In Figure 1, for example, production and exchange at quantity Q increase total welfare, since the height of the demand curve exceeds the height of the supply curve at that quantity. Triangle abc represents the aggregate increase in welfare resulting from exchanges of the good in an efficient market – in standard nomenclature, “total surplus.” Triangle abd represents the welfare increase to consumers – “consumer surplus” – and triangle bcd the welfare increase to producers – “producer surplus.”

Optimal tax theory applies this same approach to problems of taxation. Consider, for example, Harberger’s account of deadweight loss (Harberger 1964) – one of optimal tax theory’s most iconic conclusions.

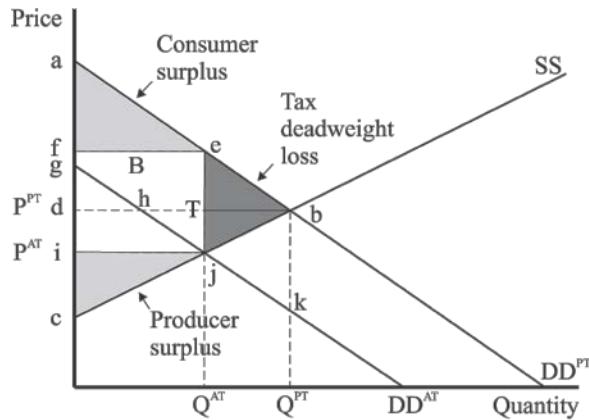


Figure 2

In Figure 2, buyer must pay tax at rate T in addition to the market price. As a result, the demand curve faced by producers is shifted downward by amount T from pre-tax demand DD^{PT} to after-tax demand DD^{AT} . The market equilibrates at after-tax price P^{AT} and quantity Q^{AT} rather than at pre-tax price P^{PT} and quantity Q^{PT} .

Prior to imposition of the tax, total surplus is represented by triangle abc, consumer surplus by triangle abd, and producer surplus by triangle bcd. After imposition of the tax, consumer surplus is represented by triangle aef and producer surplus by triangle cij. Rectangle efij represents tax revenue collected by the government, tax rate T times base B. So long as the revenues thus collected are expended on items producing equivalent welfare, this appropriation is welfare-neutral. Triangle bej, however, represents the additional amount of social welfare that would have been generated through market exchanges at price P^{PT} and quantity Q^{PT} in the absence of the tax. This foregone social welfare is the tax's "deadweight loss."

If preferences reflect welfare, tax deadweight loss represents a reduction of aggregate social welfare. If rules are morally justified only if and to the extent they maximize aggregate social welfare, taxes are therefore inherently morally suspect.

But do preferences really reflect welfare? For the most part, economists have avoided this question. One obvious problem is that no consensus definition of "welfare" exists. One might define "welfare" objectively. John Rawls, for example, postulated a class of "primary social goods" – "things that every rational man is presumed to want," those needed by anyone to live a good life and to develop their "moral powers" and other capacities. Among such goods, he identified rights and liberties, powers and opportunities, income and wealth, and the social bases of self-respect. (Rawls 1971) Unfortunately, Rawls' definition, however philosophically persuasive, cannot easily be quantified.

One might instead focus on happiness, as Jeremy Bentham and John Stuart Mill did when they first developed utilitarianism. Happiness can be measured through survey instruments; direct measurement of the relevant brain activity may also be possible. One might view happiness as evidencing the satisfaction of objective needs. One might postulate, for example, that the body knows what it needs *ex post* better than the mind knows *ex ante*. If so, happiness may constitute a measurable proxy for objective welfare. Unfortunately, happiness does not map conveniently onto standard microeconomics.

Economists therefore prefer to focus on preferences, which do so map. The premise that preferences reflect welfare exists in two versions. The first asserts that, in general, an individual's preferences constitute our best evidence of what would enhance that individual's welfare. In theory, this version is falsifiable, although falsification requires that the "welfare" of which preferences are evidence be specified, at least to some extent. The second is definitional – that welfare *is* preference satisfaction, nothing more and nothing less. Bernheim (2009) justifies use of this second version on the ground that it "leads to a rich and tractable normative framework." Tractability, not validity, is apparently dispositive.

A definitional approach, of course, eliminates any possibility of falsification. Indeed, under the definitional approach, the goal of preference satisfaction becomes an inherent part of the moral claim. As a moral claim, however, it presents serious problems – regardless of tractability.

The first is a lack of moral persuasive power. It is one thing to assert that we should pay taxes to ensure that others will have access to basic minimums – food, shelter, civil and political rights. It is another thing entirely to assert that we should pay taxes to ensure that others can satisfy their preferences, even if those preferences are for alcohol, smoking, playing the lottery, or not working. Maximization of preference satisfaction becomes a distinctly less compelling moral imperative when we are confronted with preferences that we do not share (or, perhaps, that we share but believe we should resist).

A related problem is that maximization of preference satisfaction neither fits nor justifies our current moral practices. The United States, for example, limits the earned income tax credit to working taxpayers, regardless of whether they would prefer not to work. It limits the Temporary Assistance for Needy Families program (“TANF,” more commonly known as “food stamps”) to a basket of items Congress believes will increase recipients’ welfare, again regardless of the recipients’ own beliefs and preferences. A moral theory that fails to fit or justify accepted practices is of limited utility (Dworkin 1977).

A third problem, at least for welfarists who believe in redistribution, is that preference satisfaction is not necessarily subject to the law of declining marginal utility. Those with a lot may nevertheless passionately desire more. But the law of declining marginal utility is essential to welfarist justifications for redistribution, and with it progressive taxation. Both are difficult to justify within a preference satisfaction maximization regime.

Then there is what I will call the objecting taxpayer problem. If taxpayers strongly prefer not to pay taxes, requiring them to do so is unlikely to maximize aggregate preference satisfaction. If we truly believe that maximizing preference satisfaction is a moral imperative, we should only impose taxes on those who do not strenuously object. To do otherwise would be immoral.

Finally, a definitional approach leaves open the possibility of secondary adaptation – changing preferences to fit circumstances. Instead of changing the world to satisfy pre-existing preferences, we might change those preferences to fit the world – that is, persuade everyone to be happy with what they have. All preferences would then be satisfied. Aldous Huxley’s *Brave New World* describes just such a paradise. Not all accept this as ideal.

For these and other reasons, a version of “preferences reflect welfare” that contends instead that, in general, an individual’s preferences constitute

our best evidence of what would enhance that individual's welfare seems more satisfactory. Implicit in this version is the premise that an individual's welfare is real, objective, and distinct from his or her preferences, even if in fact preferences generally reflect welfare. The last, however, is a testable factual claim – a claim which, for the most part, economists have avoided testing. Assuming that preference satisfaction is not welfare-maximizing by definition, do preferences in fact reflect welfare? It is to this question that I now turn.

Part II: The Case of Preference-Shifting

Behavioral economics acknowledges the possibility of irrationality – that is, differences between preferences and welfare – but limits its focus to distortions triggered by specific frames and heuristics. That preferences reflect welfare is assumed to be the general case; behavioral economics attempts to model limited exceptions.

I focus here instead on advertising and the welfare effects of consumerism. Unlike behavioral economics, the consumer psychology literature that services the advertising and marketing industries takes a less theoretical approach, largely indifferent to the specific heuristics or irrationalities used to trigger desired behaviors (Haugtvedt 2008). A large part of that literature describes and tests practical techniques for inducing consumers to buy more than they otherwise would, at higher prices than they otherwise would pay. Many, if not most, such techniques do not depend on the correction of informational deficits. It is therefore possible that at least some such techniques have the effect of shifting the preferences of consumers in ways that are not welfare-enhancing. If consumers' pre-advertising preferences correctly reflect consumer welfare, post-advertising preferences arguably cannot.

Gary Becker has offered the best-known response (Becker 1996). Becker hypothesizes that non-informational advertising creates intellectual capital inside each recipient's mind. This capital in turn generates imputed income. This imputed income reduces the shadow price of the good itself. For example, if a given car is worth \$300 a month to a given consumer pre-advertising, the advertising is assumed to produce a stock of intellectual capital (wholly inside the consumer's mind) that throws off \$100 per month of imputed income (again, wholly inside the consumer's mind) if he buys the car. When he is observed to pay \$400 per month for the car, this change in observed behavior does not reflect any change in preferences; the car is still only worth \$300 a month. \$100 of that \$400 is being paid for the \$100 of imputed income – not for the car. This reconstruction allows Becker to continue to assert that the consumer's preferences have not changed – and therefore allows him to continue to assume that consumers' unchanging preferences reflect welfare.

Unfortunately, Becker's theory is nonfalsifiable. Both the hypothetical stock of intellectual capital and the hypothetical flow of imputed income are unobservable; their existence, untestable. The theory also violates Occam's razor; indeed, it is Ptolemaic in its convolution. What the rest of the world calls increased desire Becker calls "imputed income" – different in kind, not merely in magnitude, from the base level of desire originally observed. Because this imputed income is not a "preference," Becker can assert that "preferences" remain unchanged: fixed, exogenous, and welfare-reflecting as they always have been. It is unlikely that any non-economist finds this plausible.

Becker's theory also leads to the rather odd conclusion that a country whose residents have been conditioned to buy things that they otherwise would not buy possesses more capital – and is therefore wealthier – than an otherwise identical country whose residents have not been so conditioned. Even economists who rely on Becker's theory, however, do not include his hypothesized intellectual capital in their national wealth accounts.

Unfortunately, measuring the size of any resulting economy-wide wedge between preferences and welfare would first require that we define "welfare" – a task that eluded both philosophers and scientists for all of recorded history. This paper will not attempt any such task; nor will it attempt to estimate the size of any such wedge. It will, however, offer two examples of possible preference-shifting to support its premise that the problem should not be ignored. The first involves changed demand for an intermediate good – consumer loans; the second, changed demand for a product as to which the objective welfare effects have been rigorously measured – a prescription drug.

Marianne Bertrand et al (Bertrand 2010) found that including a small picture of a woman in a consumer loan solicitation increased loan demand by about as much as a 25% reduction in the interest rate charged, compared with an identical solicitation without the picture. In other words, the picture permitted a 33% market equilibrium price increase. Since consumer loans are intermediate goods, it is hard to see how respondents who accepted the solicitation at the higher price could have received any commensurate welfare increase, regardless of how welfare is defined. Inclusion of the picture seems to have shifted respondents' utility and demand curves without any commensurate increase in the welfare they derived from the loans.

**the trusted way
to borrow cash**

25 September 2003 Shop 9B, Pinetown Arcade
Pinetown 3800
Tel: 031 717 8950

A special rate for you.

Congratulations! As a valued client, you are now eligible for a special interest rate on your next cash loan from [redacted]. This is a limited time offer, so please come in by 31 October 2003.

You can use this cash to buy an appliance, or for anything else you want.

Enjoy low monthly repayments with this offer! For example:

	4 Months	6 Months	12 Months
R500	R149.95	R108.28	R66.62
R1000	R299.90	R216.57	R133.23
R2000	R599.80	R433.13	R266.47
R4000	R1199.60	R866.27	R532.93

LOAN AVAILABILITY SUBJECT TO TERM & CONDITIONS

Loans available in other amounts. There are no hidden costs. What you see is what you pay.

If you borrow elsewhere you will pay R280.14 more in total on a R300.00, 4 month loan.

How to apply:
 Bring your ID book and latest paylip to your usual branch, by **31 October 2003** and ask for
Names of client, employee and Lender expressed to remain confidential.

Customer Consultant:
PS: Unfortunately, if you have already taken a loan since the date this offer was issued, you do not qualify for this offer.
 Comparison based on a standardly issued loan of R500.



FIGURE II
Example Letter 2

A second example: omeprazole, a drug sold under the brand name “Prilosec,” is used to decrease the production of gastric acid. AstraZeneca’s patent on omeprazole was due to expire in 2001. AstraZeneca therefore developed and patented an approximate enantiomer (mirror-image version) of omeprazole, esomeprazole, which it marketed as “Nexium.” Chemically, the two drugs operate the same way when ingested. Extensive studies have not shown any statistically significant differences in efficacy at equivalent doses (Drug Class Reviews 2009).

After patenting Nexium, AstraZeneca undertook a direct-to-consumer marketing campaign, built around the phrase “The Purple Pill™.” By 2009 it had succeeding in shifting consumer preferences to the point that it could charge an average U.S. price of \$190 per prescription for Nexium. Omeprazole, with medically equivalent effects at equivalent doses, commanded a price of only \$31 per prescription. The price premium for

Nexium, therefore, exceeded 500 percent. By 2009, Nexium was the second-largest-selling prescription drug in the United States, with over \$5 billion per year in sales. (Both omeprazole and esomeprazole have since been reclassified as over-the-counter drugs. As a result, Nexium's price premium has largely disappeared.)

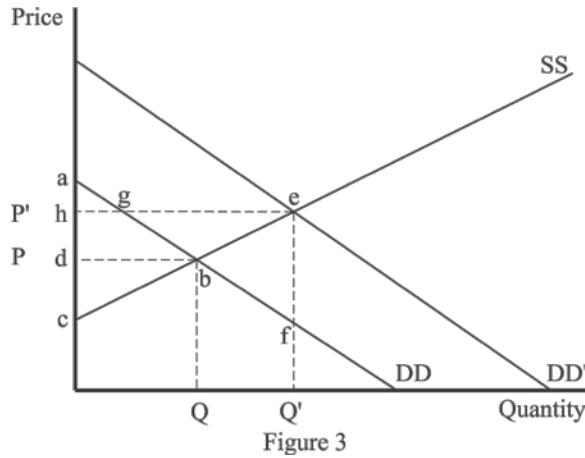
Nexium's \$159 per prescription premium was arguably evidence of preference-shifting: an upward shift in observed utility and demand curves without any corresponding increase in consumer welfare. The advertising theme used to effect this shift – “The Purple Pill™” – was largely devoid of informational content. With respect to Nexium, at least, Becker's accumulation-of-intellectual-capital explanation has little explanatory power. One might reasonably doubt that consumers received \$159 per prescription of additional welfare by reason of their decisions to purchase Nexium rather than omeprazole.

Neither of the foregoing examples nor any other study or anecdote establishes the existence of a general wedge between preferences and welfare. For present purposes, it is sufficient that economically significant preference-shifting is plausible. If preference-shifting can be shown to change standard results materially, the need for tools that will allow us to estimate the size of the wedge between preferences and welfare becomes more pressing. Does “Super Size Me!” shift demand upward without a commensurate increase in objective consumer welfare? Such questions are empirical and largely beyond the capacity of currently available econometric tools. This paper asks instead whether the development of such tools might be worth the effort – specifically, whether differences between observed choices and welfare, if they do exist, require adjustments to important parts of the optimal tax canon.

Part III: Modeling Preference-Shifting

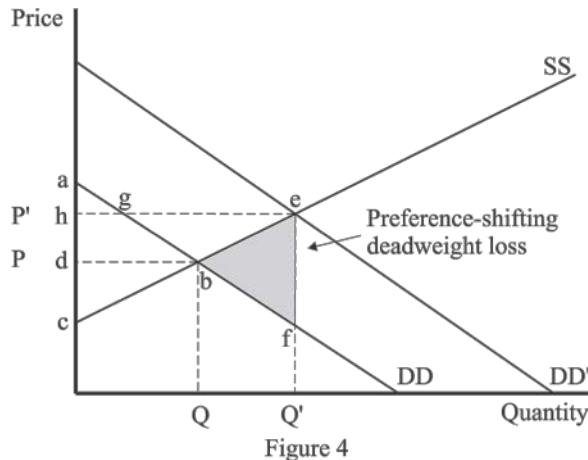
Recall that the standard model assumes that the height of the demand curve at any quantity equals the marginal social benefit of consumption of that quantity and that the height of the supply curve equals the marginal social cost of production of that quantity, all costs and benefits having been internalized. Under these assumptions, total surplus provides a measure of the net increase in social welfare by reason of market exchanges at the equilibrium price and quantity. If so, then in the absence of market failure, supply and demand adjust automatically to maximize social welfare. Critically, this conclusion depends on the assumption that the choices embodied in observed demand reflect welfare.

The effects of preference-shifting can be modeled by postulating two demand curves: a non-preference-shifted, welfare-reflecting curve and a preference-shifted curve. In the absence of intervention, supply and demand will equilibrate at the intersection of supply and preference-shifted demand.



In Figure 3, DD is the non-preference-shifted, welfare-reflecting demand curve and DD' the preference-shifted curve. In a non-preference-shifted world, the market would equilibrate at price P and quantity Q, with the same welfare results as in Figure 1. Because preferences have been shifted, however, the market equilibrates at price P' and quantity Q'.

The marginal social benefit of consumption of quantity Q' is given by the height of the non-preference-shifted demand curve DD at Q', the marginal social cost by the height of the supply curve at Q'. For all quantities between Q and Q', marginal social cost exceeds marginal social benefit. Total social loss by reason of preference-shifting is therefore represented by triangle bef in Figure 4 below, the “preference-shifting deadweight loss.” (Costs incurred by producers to effect the preference shift also constitute social loss; this paper ignores such costs.)



More generally:

$$\text{Preference-shifting deadweight loss} = \int_Q^{Q'} d(q) - s(q),$$

Where:

Q = quantity exchanged at the welfare-maximizing equilibrium

Q' = quantity exchanged at the observed equilibrium

$d(q)$ = marginal social benefit as a function of quantity consumed, and

$s(q)$ = marginal social cost as a function of quantity produced.

If $Q' > Q$ and $\frac{ds}{dq} > \frac{dd}{dq}$ between Q and Q' , preference-shifting will produce net deadweight loss.

In figure 5 below, producer surplus is represented by quadrilateral cbgh. In addition, an amount represented by triangle beg is redistributed from consumers to producers. This is not “surplus,” because it does not represent an increase in social welfare by reason of market exchanges; it is pure redistribution. Total gain to producers from exchanges in a preference-shifted market, represented by triangle ceh, is the sum of producer surplus and this preference-shifting redistribution. Or:

$$\text{Producer gain from market exchanges} = \int_0^{Q'} P' - s(q)$$

$$\text{Producer preference-shifting profits} = Q * (P' - P) + \int_Q^{Q'} P' - s(q)$$

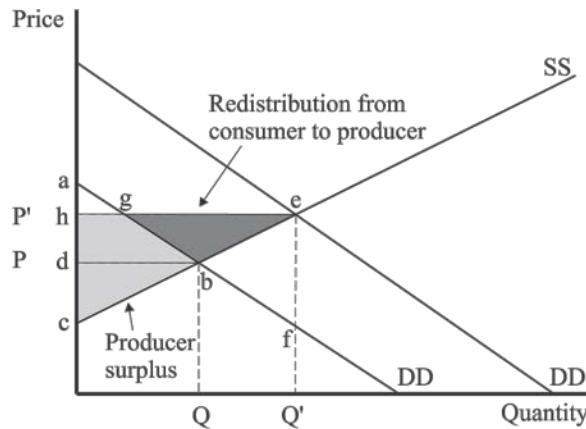


Figure 5

The computation of consumer gain or loss in a preference-shifted market begins with consumer surplus, represented by triangle agh. From this, however, must be subtracted both the preference-shifting redistribution from consumers to producers, represented by triangle beg,

and the preference-shifting deadweight loss, represented by triangle bef.
Or:

$$\text{Consumer gain from market exchanges} = \int_0^{Q'} d(q) - P'$$

Note that, as in Figure 6, the result may be negative. If so, market exchanges of the preference-shifted good reduce consumer welfare.

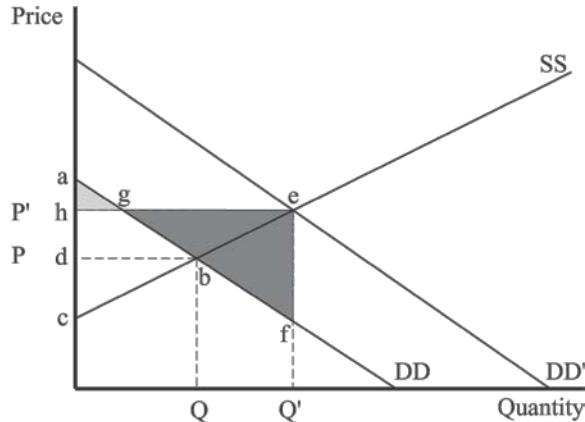


Figure 6

Finally, in Figure 7 net social benefit or loss by reason of market exchanges of the good is represented by total surplus, triangle abc, less preference-shifting deadweight loss, triangle bef. More generally:

$$\text{Social gain from market exchanges} = \int_0^{Q'} d(q) - s(q)$$

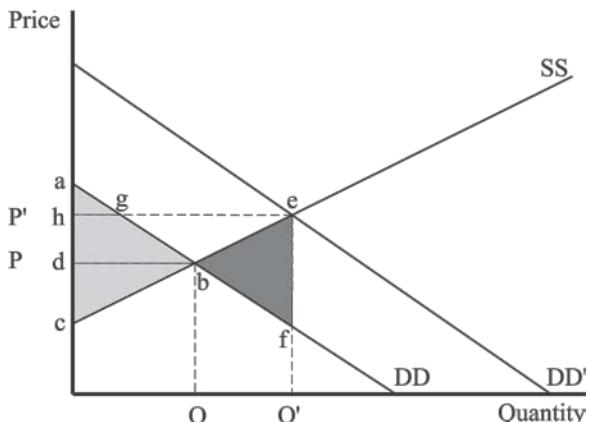


Figure 7

There is no inherent upper limit to preference-shifting deadweight loss. If preference-shifting deadweight loss is sufficiently large, as in Figure 8, the net social welfare consequences of exchanges in a preference-shifted market can be negative. (Recall that the 2009 price premium for Nexium was over 500 percent.)

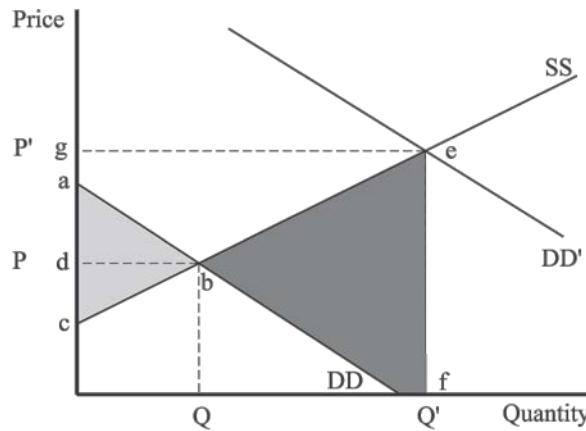


Figure 8

The effects of preference-shifting can therefore be summarized as follows.

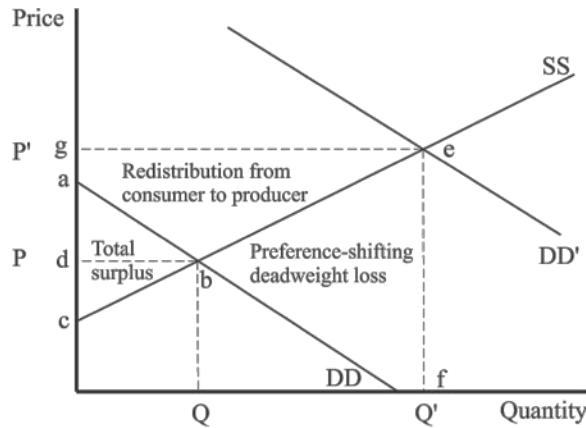


Figure 9

Total surplus is represented by the area between the supply curve and the welfare-reflecting demand curve to the left of quantity Q .

$$\text{Total surplus} = \int_0^Q d(q) - s(q)$$

Preference-shifting deadweight loss is represented by the area between the supply curve and the welfare-reflecting demand curve between quantities Q and Q' . No additional surplus is generated by exchanges of quantities above Q ; such exchanges instead produce social loss.

$$\text{Preference-shifting deadweight loss} = \int_Q^{Q'} d(q) - s(q)$$

Finally, preference-shifting redistribution from consumers to producers is represented by the area above the supply curve and the welfare-reflecting demand curve bounded by price P' .

$$\text{Preference-shifting redistribution} = \int_0^Q P' - d(q) + \int_Q^{Q'} P' - s(q)$$

The foregoing results obtain whenever preferences do not correctly reflect objective welfare, regardless of whether intentional preference-shifting has occurred, so long as observed demand exceeds welfare-reflecting demand. These results obtain even if the activities that shift consumer preferences also enhance welfare (for example, by conveying useful information or shifting shadow prices) and thereby shift the welfare-reflecting demand curve upward, so long as observed demand is increased more than welfare is increased (that is, so long as the objective welfare increase is not commensurate).

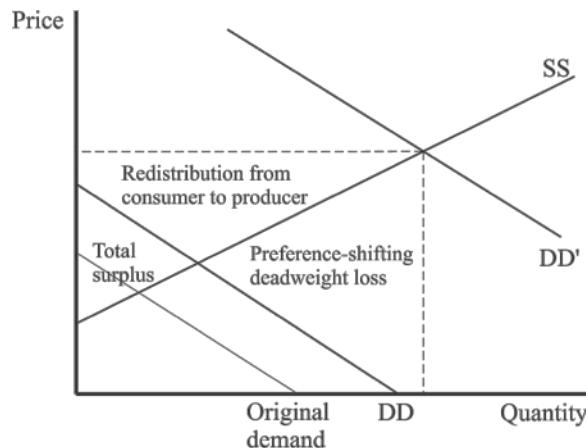


Figure 10

In Figure 10, producers' preference-shifting activities shift demand in ways that increase consumer welfare, but not commensurately with the increase in observed demand. The new welfare-reflecting demand curve is DD , the preference-shifted demand curve DD' . Total surplus, preference-

shifting deadweight loss, and redistribution from consumers to producers are then the same as in Figure 9.

Part IV: Effects of Tax on Deadweight Loss

In the standard account, a commodity tax can be modeled either as a downward shift in demand or an upward shift in supply, depending on the party on which it is nominally imposed, in each case by the amount of the tax. The incidence of the tax will depend on the relative elasticities of supply and demand, not on the government's choice of nominal taxpayer.

Recall that in Figure 2, triangle $b ej$ represents the additional amount of social welfare that would have been generated through market exchanges at price P^{PT} and quantity Q^{PT} in the absence of the tax. This foregone social welfare is the "tax deadweight loss."

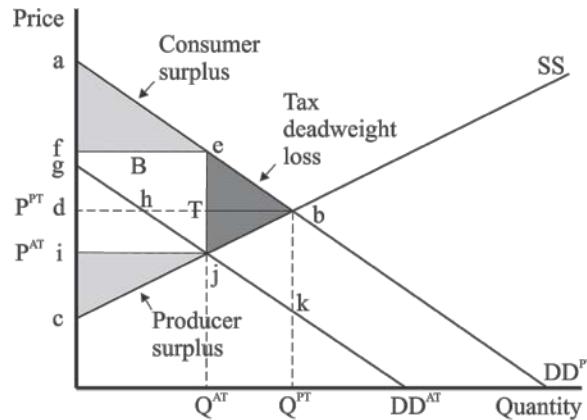


Figure 2

We can now model imposition of a tax on a good for which demand is preference-shifted. Prior to imposition of the tax,

$$\text{Total surplus} = \int_0^Q d(q) - s(q), \text{ and}$$

$$\text{Preference-shifting deadweight loss} = \int_Q^{Q'} d(q) - s(q),$$

Where:

Q = quantity exchanged at the welfare-maximizing equilibrium

Q' = quantity exchanged at the observed pre-tax equilibrium

$d(q)$ = marginal social benefit as a function of quantity consumed, and

$s(q)$ = marginal social cost as a function of quantity produced.

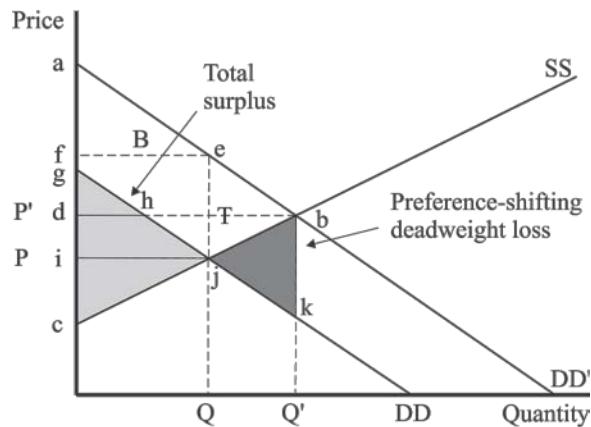


Figure 11

Consider tax at rate T in an amount equal to the preference-shift. The tax forces the market to reequilibrate at price P and quantity Q – the same equilibrium price and quantity the market would have reached in the absence of preference-shifting. Buyers pay price P plus tax at rate T, and at that effective price only demand quantity Q.

Two consequences ensue. First, because no exchanges occur in excess of quantity Q, preference-shifting deadweight loss is eliminated. Second, market exchanges at price P' and quantity Q' would not have generated any additional social welfare; the forced reequilibration at price P and quantity Q therefore does not result in any tax deadweight loss. As a result, imposition of tax at a rate less than or equal to the preference-shift increases aggregate social welfare.

Part V: Incidence of the Tax

If the tax rate equals the preference-shift, the resulting increase in aggregate social welfare will equal the preference-shifting deadweight loss that would have been incurred in the absence of the tax.

$$\text{Total welfare gain from tax} = - \int_Q^{Q'} d(q) - s(q)$$

This welfare gain is allocated among consumers, producers, and the government as follows.

The government receives the amount of the tax.

$$\text{Government receipts} = Q * T$$

Producer welfare is reduced by the amount of producers' pre-tax preference-shifting profits, represented by quadrilateral bdij in Figure 12 below. This consists of (1) the portion of consumer surplus appropriated

by producers by reason of the preference-shift, represented by quadrilateral dhij, plus (2) the preference-shifting redistribution from consumers to producers, represented by triangle bhj. As a result, producer welfare gains by reason of market exchanges at the after-tax equilibrium exactly equal the producer surplus that would have been generated by market exchanges in the absence of both tax and preference-shifting.

$$\text{Producer incidence} = - Q * (P' - P) - \int_Q^{Q'} P' - s(q)$$

If preference-shifting is a form of market failure, from producers' perspective the tax exactly corrects the market failure.

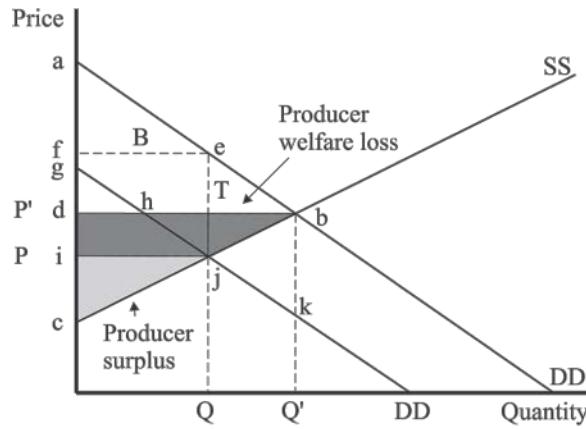


Figure 12

Net consumer welfare gain or loss by reason of the tax, depicted in Figure 13 below, equals (1) the preference-shifting deadweight loss plus the preference-shifting redistribution from consumers to producers in excess of quantity Q, reduced by (2) the portion of the tax in excess of the change in price (P' minus P).

$$\text{Consumer incidence} = [\int_Q^{Q'} P' - d(q)] - [Q * (P + T - P')]$$

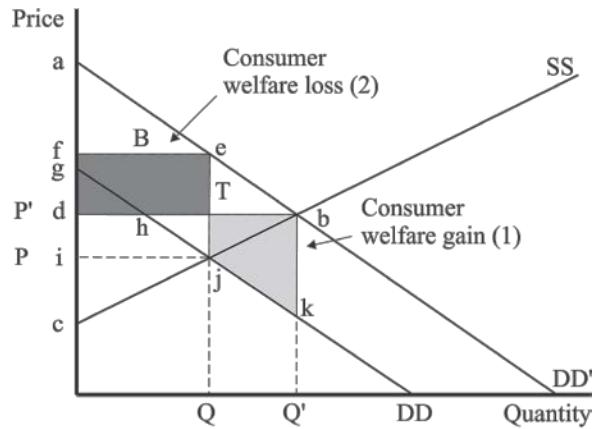


Figure 13

If (1) is greater than (2),

$$\int_Q^{Q'} P' - d(q) > Q * (P + T - P'),$$

the tax will result in net welfare gain to consumers in the amount of the difference. In such event, producers will bear the entire burden of the tax, but only in the amount of their preference-shifting profits. If (2) is greater than (1),

$$\int_Q^{Q'} P' - d(q) < Q * (P + T - P'),$$

consumers will bear a portion of the tax equal to such excess. A portion of the tax equal to the preference-shifting deadweight loss will be recovered through elimination of that loss and will not be borne by any human being.

Part VI: Conclusions and Further Work

Optimal tax theory is currently based on the factual assumption that preferences reflect welfare. In practice, this factual assumption is neither tested nor questioned – it is effectively immune from the norm of falsifiability. The pervasiveness of advertising in modern consumer economies, however, makes optimal tax theory's failure to test this factual assumption deeply problematic. The available evidence, such as it is, appears to confirm that significant wedges may exist between preferences and welfare – indeed, that businesses may intentionally create such wedges.

This paper postulates the possibility of “preference-shifting” – where businesses shift the preferences of potential consumers in ways that are not commensurately welfare-enhancing and then satisfy the resulting shifted

preferences – and suggests a simple way to incorporate preference-shifting into the standard model. It then demonstrates that, in the case of a commodity tax, imposition of tax at a rate less than or equal to the preference-shift increases aggregate social welfare. Imposition of tax in an amount equal to the preference-shift will return producer surplus to the amount that would have accrued to producers in the absence of preference-shifting at the welfare-maximizing equilibrium. Producers will bear the burden of such a tax to the extent of their pre-tax preference-shifting profits. Consumers will bear the burden of such a tax to the extent such burden is not borne by producers, but will also benefit from the elimination of preference-shifting deadweight loss. A portion of such tax equal to the preference-shifting deadweight loss will be recovered through elimination of that loss and will not be borne by any human being.

In the standard account, once it has been shown that commodity taxation produces deadweight loss, the result is generally extended to income and other forms of taxation without further demonstration. If preference-shifting were to operate randomly with respect to income, it would be necessary to customize commodity taxes to reflect such random preference-shifting. It is plausible, however, that preference-shifting is more likely with respect to non-essentials and therefore increases with income. If so, progressive income taxation may reduce income-correlated preference-shifting deadweight loss and thereby increase aggregate welfare.

Standard accounts of the welfare effect of labor tax rates (*e.g.*, Mirrlees 1971) assume that preferences reflect welfare. Such accounts typically begin at the individual utility curve level. Demand curves, of course, are simply aggregations of individual utility curves. The model offered in this paper assumes two sets of utility curves – a set of non-preference-shifted, welfare-reflecting utility curves and a set of preference-shifted utility curves. If a wedge exists between these two sets, Mirrlees' analysis and its progeny require modification.

A more immediate and perhaps less theoretically contestable implementation of the commodity tax results might be to limit retail sales and value added tax exemptions to non-advertised goods. This would merely require that we assume that, in general, preference-shifting by reason of advertising equals or exceeds the retail sale or value added tax rate. Since such rates are normally small relative to price, such an assumption is plausible. The Atkinson-Steiglitz theorem, which suggests that governments should abstain from differentiated commodity taxation if non-linear income taxation is an option, does not apply in this context because it too is based on the assumption that preferences are fixed and exogenous and reflect welfare (Atkinson-Steiglitz 1976).

In any event, the fact that preference-shifting implies results very different from those of the standard model makes more urgent the need for tools to measure the wedge, if any, between observed preferences and

welfare. This may, in turn, require further definition of “welfare” – a task that economics has heretofore largely avoided. The standard account asserts that whatever “welfare” means, if observed preferences reflect welfare, taxes result in tax dead-weight loss. This paper has shown that if observed preferences do not reflect welfare, taxes do not necessarily result in dead-weight loss – indeed, that taxes may enhance “welfare,” again regardless of what that term means. To move from theory to measurement, however, may require further definition.

“Capitalism,” said Andrew Carnegie, “is about turning luxuries into necessities.” Preference-shifting may be necessary to capitalism. Nothing in this paper is intended to suggest that either is undesirable. Its conclusion is purely technical: that if economically significant preference-shifting exists, the optimal tax canon requires adjustment. If important consequences of the standard account depend on an assumption – that preferences reflect welfare – possibly false in economically significant ways, it would be inappropriate to continue to assert the results of the standard model without appropriate caveat.

References

- Atkinson, A. B.; Stieglitz, J. E., The Design of Tax Structure: Direct versus Indirect Taxation, *Journal of Public Economics* 6 (1-2): 55–75 (1976).
- Becker, Gary, *Accounting for Tastes* (Harvard Univ. Press 1996).
- Bernheim, B. Douglas, Behavioral Welfare Economics, 7(2-3) *J. Eur. Econ. Ass'n* 267-319 (2009).
- Bertrand, Marianne, et al, What's Advertising Content Worth? Evidence From a Consumer Credit Marketing Field Experiment, *Quarterly Journal of Economics* (2010) 125 (1): 263-306.
- Drug Class Reviews, Proton Pump Inhibitors, Final Report (Update 5, May 2009).
- Dworkin, Ronald, *Taking Rights Seriously* (1977)
- Harberger, Arnold C., The Measurement of Waste, 54 *Am. Econ. Rev.* 58 (1964).
- Haugtvedt, Curtis P., et al, eds., *Handbook of Consumer Psychology* (Psychology Press 2008)
- Mirrlees, J. A., An Exploration in the Theory of Optimum Income Taxation, 38 *Rev. Econ. Stud.* 175-208 (1971).
- Rawls, John, *A Theory of Justice* (1971).
- Slemrod, Joel, Optimal Taxation and Optimal Tax Systems, 4 *J. Econ. Persp.* 157 (1990).